Using sentiment analysis in tourism research: A systematic, bibliometric, and integrative review
Manosso, Franciele Cristina; Domareski Ruiz, Thays Cristina

Empfohlene Zitierung / Suggested Citation:

Nutzungsbedingungen:
Dieser Text wird unter einer CC BY-NC-ND Lizenz (Namensnennung-Nicht-kommerziell-Keine Bearbeitung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
https://creativecommons.org/licenses/by-nc-nd/4.0/deed.de

Terms of use:
This document is made available under a CC BY-NC-ND Licence (Attribution-Non Comercial-NoDerivatives). For more Information see:
https://creativecommons.org/licenses/by-nc-nd/4.0

Diese Version ist zitierbar unter / This version is citable under:
https://nbn-resolving.org/urn:nbn:de:0168-ssoar-75146-7
Using sentiment analysis in tourism research: A systematic, bibliometric, and integrative review

Franciele Cristina Manosso
Federal University of Paraná, Brazil

Thays Cristina Domareski Ruiz
Federal University of Paraná, Brazil

Abstract:

Purpose: Sentiment analysis is built from the information provided through text (reviews) to help understand the social sentiment toward their brand, product, or service. The main purpose of this paper is to draw an overview of the topics and the use of the sentiment analysis approach in tourism research.

Methods: The study is a bibliometric analysis (VOSviewer), with a systematic and integrative review. The search occurred in March 2021 (Scopus) applying the search terms "sentiment analysis" and "tourism" in the title, abstract, or keywords, resulting in a final sample of 111 papers.

Results: This analysis pointed out that China (35) and the United States (24) are the leading countries studying sentiment analysis with tourism. The first paper using sentiment analysis was published in 2012; there is a growing interest in this topic, presenting qualitative and quantitative approaches. The main results present four clusters to understand this subject. Cluster 1 discusses sentiment analysis and its application in tourism research, searching how online reviews can impact decision-making. Cluster 2 examines the resources used to make sentiment analysis, such as social media. Cluster 3 argues about methodological approaches in sentiment analysis and tourism, such as deep learning and sentiment classification, to understand the user-generated content. Cluster 4 highlights questions relating to the internet and tourism.

Implications: The use of sentiment analysis in tourism research shows that government and entrepreneurship can draw and enhance communication strategies, reduce cost, and time, and mainly contribute to the decision-making process and understand consumer behavior.

Keywords: Sentiment analysis, tourism, bibliometrics, systematic review, integrative review, Vosviewer

JEL Classification: L83, C38, Z30

Biographical note: Franciele Cristina Manosso is a Doctoral Researcher in Business Administration with emphasis on consumer behavior, at Federal University of Paraná, Brazil. E-mail: francimanosso@gmail.com. Thays Cristina Domareski Ruiz is a Professor at the Department and Master’s Program of Tourism, Federal University of Paraná, Brazil. E-mail: thaysruiz@ufpr.br.

1 INTRODUCTION

In recent years, consumers have changed the way they use social media, which means that the purpose of visiting and using these websites became more interactive with people sharing information about their daily lives’ experiences and the products and services they consume (Cambria, Schuller, Xia & Havasi, 2013; Kim, Park, Yun & Yun, 2017). This context emerges because digital technology and social media have a substantial impact on the way people share information and opinions (Sheth, 2020), so it can be considered as an important tool to understand public expression about a relevant event or an experience that has been lived in a hotel, for example (Deng, Gao, Wang & Zhang, 2020).

Social media and a large amount of information shared in these platforms impacts, directly, the hospitality domain (García-Pablos, Cuadros & Linaza, 2016) as a result of the necessity that tourists have to seek facts about a destination and the experiences other people lived there (Yan, Zhou & Wu., 2018). The tourists became aware of the thoughts and emotions that other consumers have, and this happens because users can create posts in real-time on business pages that can be monitored and turned into actionable knowledge (Nechoud, Ghidouche & Seraphin 2021; Amanatidis, Mylona, Mamalis & Kamenidou, 2020; Maio, Fenza, Loia & Orciuoli, 2017).

It is important to highlight those travelers, as mentioned, go beyond hotel and airline tickets reservation but also exchange information and detailed descriptions of their pleasant and unpleasant travel experiences through social media and online review websites (Halkiopoulos, Antonopoulos, Papadopoulos, Giannoukou & Gkintoni, 2020; Kim et al., 2017). In this case, online reviews help consumers make more informed decisions and help them to choose the product.
or service that better fits their necessity or desire (Cheung & Thadani, 2012; Liu & Park, 2015). With the advent of social networks and web communities, a large amount of information is more attractive as a source of data for opinion mining and sentiment analysis (SA) (Cambria et al., 2013). Therefore, SA and text mining have been used as a way to get insights from online reviews in the hospitality sector (Hu & Chen, 2016). SA involves the deciphering of the opinions contained in the written or spoken word, through IT processes, to extract subjective information; opinions and feelings in regard to analytical sources are observed (Pang & Lee, 2008).

There has been growing interest in SA in detecting valence and emotions in the texts obtained from social media platforms (Mohammad, 2017). Several authors have used this technique to comprehend and measure sentiment level on airline services (Liu & Tan, 2014), hotels (Mankad, Han, Goh & Gavirneni, 2016; García-Pablos et al., 2016; Hu & Chen, 2016; Moro, Ramos, Esmerado & Jalali, 2019; Valdivia, Hrabova, Chaturvedi, Luzón, Troiano, Cambria & Herrera, 2019) and destination experiences (Kim et al., 2017; Zheng, Luo, Sun, Zhang & Chen, 2018; Marine-Roig, 2019), where the sentiment is located on the positive-negative scale equated with the evaluation-potency-activity model (EPA) (Pang & Lee, 2008; Liu, 2017).

Besides this context, sentiment has been discussed as a relevant variable in marketing, because it can be used as a critical indicator regarding consumer behavior and their feelings about a purchase and an experience that was lived before (Mishkin, 1978; Gaski & Etzel, 1986; Thropp, 1992; Carroll, Fuhrer & Wilcox, 1994; Kim et al., 2017). Hence, it is important to mention that online reviews have become an instrument for consumers to make better and conscious decisions (Yan et al., 2018), and it's a way to help both businesses and consumers because one can draw more relevant experiences.

In this case, applying SA and opinion mining in studies that have the objective to understand the consumers’ feelings and opinions is important to deliver better and competitive experiences, products, and services (Hussain & Cambria, 2018). The main purpose of this paper is to draw an overview of the topics and the use of the SA approach in tourism research, which raises the following question: how has tourism research used the SA approach to understand the information shared by the consumer in social media?

The paper is structured as follows. First, a literature review is provided on SA in tourism research. Second, the research method employed for this study is explained. Next, the research findings are presented and discussed. Finally, the study conclusions. This article provides an overview of the SA and highlights future directions for tourism research.

2 BACKGROUND LITERATURE

In the last decades, researchers and even entrepreneurs have seen an explosion of text data generated by consumers in many ways, such as text messages, reviews, tweets, emails, posts, and blogs (Humphreys & Wang, 2018). People nowadays share, use, and search for information online regularly (Hemmatian & Sohrabi, 2019); more importantly, social media has become a tool that can be used for those purposes. In this case, it is possible to read different opinions, and, inside this context, many sentiments, evaluations, emotions, appraisals, and attitudes emerge from this content (Liu, 2017). So, the internet and social media have transformed our communication; web text is becoming one of the most important channels for people to express their opinions, mental state and communicate with each other (Batrinca & Treleaven, 2015).

Hence, to understand consumer opinions, a new research field called SA emerges (Serrano-Guerrero, Olivas, Romero & Herrera-Viedma, 2015), which is the computational study of people’s opinions, appraisals, attitudes, and emotions toward entities, individuals, issues, events, topics, and their attributes (Prabowo & Thewall, 2009; Liu & Zhang, 2012; Kirilenko et al., 2018). It is important to point out that the focus of SA is to extract a sentiment expressed in a document toward a certain aspect based on the subjectivity and the linguistic characteristics of the words within an unstructured text (García, Gaines & Linaza, 2012).

The foundation of SA is based on the assumption that the arrangement of the theoretical and methodological approach of social networks is the recognition that the causal mechanism is what the consumers do, think, or feel, lies in the patterns of relationships between the actors themselves, caught in a given context and reference to an identifiable temporal frame (Micera & Crispino, 2017).

As mentioned, SA searches to comprehend the emotions and sentiments in the opinions shared by different consumers; thus, it is important to differentiate these constructs. First, emotion is an intuitive feeling that is triggered by situational cues. Meanwhile, sentiment is an organized feeling or accumulated emotions, and it is highly socialized and developed from thought rather than being instinctive emotions (Liu et al., 2019).

Figure 1: Sentiment classification techniques

Source: Authors (2021) based on Serrano-Guerrero et al. (2015, p. 21).

These sentiments can be categorized into positive and negative; or into an n-point scale, e.g., very good, good, satisfactory, bad, very bad (Prabowo & Thewall, 2009). Several emotions and affective states can be assigned positive or negative valence as well; for example, joy is considered as carrying positive valence and, thus, indicates positive sentiment while anger is indicative of negative sentiment. The intensity of the sentiment can be measured by how far from a neutral point on the positive-negative dimension a
concept is located (Kirilenko et al., 2018). Consequently, Serrano-Guerrero et al. (2015) present two approaches to perform SA, see Figure 1.

Hence, there is the Machine Learning Approach that is divided into supervised and unsupervised techniques, the first one is relevant when there is a defined corpus to classify, and it is possible to use Support Vector Machine (SVM); Naive Bayes among others. However, in the unsupervised technique, there is no possibility of classification, but it is possible to make a hybrid approach and use both together to broadly understand sentiments, feelings, and emotions in several texts (Ye, Zhang & Law, 2009; Ganesan, Zhai & Viegas, 2012). In addition, the Lexicon-based Approach employs, mainly, the sentiment lexicon, which means that dictionaries and other structures can be used to classify the emotions that appeared in the analyzed text (i.e., reviews) (Wilson, Hoffmann, Somasundaran, Kessler, Wiebe, Choi, Cardie, Riloff & Patwardhan, 2005; Serrano-Guerrero et al., 2015).

According to Drus and Khalid (2019), papers that discuss SA on social media employ either Lexicon-based approach, Machine Learning, and, also, a mix of both methods. Further discussions about how SA can be useful and, even, which techniques are indispensable to run several kinds of research have been gained a lot of prominences, which means that it is possible to apply several resources to analyze SA, such as General Inquirer Lexicon; Sentiment Lexicon; MPQA Subjectivity Lexicon; SentiWordNet; Emotion Lexicon; Financial Sentiment Lexicons, among others (Cambria, Das, Badypadhyay & Feraco, 2017), qualitative and quantitative alternatives.

In the tourism context, it is possible to mention that, nowadays, people use the internet to search for different kinds of information, which means that opinions and sentiments expressed in a review became an important tool for the tourist decision-making process (Yan, Zhou & Wu, 2015). Consequently, tourism user-generated content (UGC) became an important instrument for comprehending consumer behavior, drawing new services based on previous experiences, and delineating marketing campaigns (Alamoudi & Alghamdi, 2021). In sum, Figure 02 highlights the main characteristics of tourists’ sentiments and Tourism UGC.

Figure 2: Tourists’ sentiments and tourism UGC

Thus, SA seeks to deliver a broad comprehension of the elements presented in Figure 02, which means that the techniques search for ways to deliver answers about emotions, and opinions that consumers have on services, products, and past experiences. In this case, besides the debate about what is SA and these applications, the present research highlights that some tourism scholars have placed tourist sentiment into two major categories, positive and negative emotions, and further divided them into several basic types of emotions, such as happiness, love, fear, anger, sadness, and regret, and thus investigated tourist sentiment factors in addition to their possible implications (Mitas, Yarmal, & Chick, 2012).

In this case, it is essential to highlight that tourism researchers have been using SA to understand the activity from new perspectives. Some themes are relevant in the tourism and SA research, such as tourism recommendation (Luo, He, Mou, Wang & Liu, 2021a; Liang, Pan, Gu, Guan & Tsai, 2021); COVID-19 (Sontayasara, Jarjapanponboon, Promjuni, Seelpipat, Saengtabtim, Tang & Leelawat, 2021); geolocation (Paolanti, Mancini, Frontoni, Felicetti, Marinelli, Marcheggioni & Perdicca, 2021); gastronomy (Yu & Zhang, 2020); cruise (Wu, Dong & Xiong, 2020); cultural tourism (Liang et al., 2021); hotels (Ray, Garain & Sokar; 2021; Hu & Chen, 2016); Airbnb (Serrano, Ariza-Montes, Nader, Sianes & Law, 2020; Cheng & Jin, 2019); and others that will be discussed in the main findings in the present research.

3 RESEARCH METHODOLOGY

The present research employs a bibliometric analysis with the support of the software VOSviewer version 1.6.16 (Van Eck & Waltman, 2010). Bibliometric studies are represented by their use of statistics to analyze the content of academic literature in a specific field over a given period. They have become an increasingly significant issue in tourism studies (Hall, 2011).

Bibliometric analysis is a quantitative research methodology that uses data from publications, sources, years, countries, and citations by providing indicators of research production in a determined area over time, allowing scholars to map this data. Scholars can identify patterns through the databases, choosing single journals or multiple publication sources or specific keywords (Ellegard & Wallin, 2015). Consequently, there are some benefits for the researchers that use this type of methodology to construe their theoretical backgrounds, such as the systematization of specific information (i.e. articles, journals, researchers, institutions, and countries); comprehension about a particular field and the networks created around the subject; and impact the research rankings among others benefits that make the bibliometric analysis a reliable methodology in different research areas (Osareh, 1996; Sigala & Christou, 2006; Ellegard & Wallin, 2015). Additionally, a systematic review was made to discuss how researchers are employing the SA methodological approach. According to Clarke and Horton (2001), a systematic review is a type of investigation focused on the reunion, critical evaluation, and synthesis of the primary research results. It aims to extract information from the articles based on the needs of the research, such as locations in which the studies were conducted, the authors, and the methods used (Cheng; Edwards; Darcy; Redfern, 2016). It is a type of review in which there is a search for relevant studies on a specific topic,
and those identified are then examined and synthesized. Its benefits become salient when a researcher wants to present general knowledge, historical perspective, identifying what is missing in a particular topic, and justify the subject relevance for future studies (Aromataris & Pearson, 2014). Finally, an integrative review was engaged to discuss the main results of the reunited papers (Broome, 2000). This kind of methodological approach helps the researchers comprehend how the topic has been debated in different areas. According to Souza, Silva, and Carvalho (2010), an integrative review has as a main purpose to provide a knowledge synthesis, and also a broad comprehension about the applicability of the results in the studies compiled for the research. Thus, an integrative review search to comprehend the results of different studies, which means that a lot of methodologies can be evaluated, and their results can be analyzed and discussed from an integrative perspective because the main core of this methodological approach is to investigate the research results and conclusions (Whittemore & Knafl, 2005; Nella & Christou, 2021).

Therefore, the present research utilizes these three approaches, mainly because it is relevant to understand a topic in a broad manner, which means that every methodological tool has an important role in the discussion about SA and its use in the tourism context. Additionally, for better comprehension, Figure 3 presents a flowchart with the research steps to help the readers and other researchers to understand the research process.

Figure 3: Research flowchart

As expected, the two words the most repeated are sentiment analysis (81) and tourism (30). Hence, as observed in Figure 4, four main clusters emerged from the results, and they include a range of papers that discuss relevant subjects for each cluster.

Cluster 01 discusses in many ways how SA can be used to draw recommendations systems that help consumers and entrepreneurs make better decisions (Santamaria-Granado, Mendoza-Moreno & Ramirez-Gonzalez, 2021; Abassi-Moud, Vahdat-Nejad & Sadri, 2021; Vissuwasan, Gladis, Kalaiselvi, Ananya & Kritika, 2020; Guerreiro & Rita, 2020; Shao, Tang & Bao, 2019); how hotels and restaurants have been used SA (Kim, Lee, Choi & Kim, 2021; Marcolin et al., 2021; Ray et al., 2021; An, Ma, Du, Xiang & Fan, 2020; Fuentes-Moraleda, Diaz-Perez, Orea-Giner, Muñoz-Mazón & Villacé-Moliner, 2020; Yu & Zhang, 2020; Park, Kang, Choi & Han, 2020; Zapata, Murga, Raymundo, Domínguez, Moguerza & Alavarez, 2019; Moro et al., 2019), and at last the papers discuss SA and the role of online reviews to build up new perspectives to understand consumer behavior, for example (Sontayasara et al., 2021; Gour, Aggarwall & Erdem, 2021; Luo et al., 2021a; Hao, Fu, Hsu, Li & Chen, 2020; Ainin, Feizollah, Anuar & Addullah, 2020; Alosaimi, Alharthi, Alghamdi, Alsubait & Alquarashi, 2020).

Table 1: Main topic and authors - Cluster 01

<table>
<thead>
<tr>
<th>Cluster 01</th>
<th>Discuss how SA has been used as a tool for entrepreneurs and consumers to make more conscious and better decisions.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kim et al., 2021; Santamaria-Granado et al., 2021; Abassi-Moud et al., 2021; Marcolin et al., 2021; Ray et al., 2021; Sontayasara et al., 2021; Gour et al., 2021; Luo et al., 2021a; Hao et al., 2020; Ainin et al., 2020; Alsubait et al., 2020; An et al., 2019; Fuentes-Moraleda et al., 2020; Yu &amp; Zhang, 2020; Park et al., 2020; Vissuwasan et al., 2020; Guerreiro &amp; Rita, 2020; Bao et al., 2019; Zapata et al., 2019; More et al., 2019.</td>
<td></td>
</tr>
</tbody>
</table>

Source: Authors (2021)
Cluster 2 also includes debates about how tourism management can use social media, big data, and text mining to understand tourist behavior and presents more competitive and relevant strategies in front of the new market. So, some articles discuss the management perspective in the national parks (Mangachena & Pickering, 2021); cultural tourism (Liang et al., 2021), and, also, in a general manner, which means that the articles include debates about destination image (Jiang, Chen, Eichelberger, Ma & Pikkemaat, 2021; Nowacki & Niezgoda, 2020; Marine-Roig, 2019); tourism management itself (Baumgartner & Zhu, 2020; Collado et al., 2020) and also how satisfaction (Chen et al., 2020) and perception (Vallone & Veglio, 2019) impacts tourist behavior and how SA is relevant. In Cluster 2, some papers make reference to TripAdvisor as a way to comprehend consumer behavior through online reviews and big data (Lee et al., 2020; Sangkaew & Zhu, 2020; Valdivia et al., 2020).

Table 2: Main topic and authors - Cluster 02

<table>
<thead>
<tr>
<th>Cluster 02</th>
<th>Discuss ways tourism management can employ SA and big data to understand consumer behavior and draw new products, services, and experiences.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Francisco e Cristina Manosso &amp; Thays Cristina Domareski Ruiz</td>
<td>Mangachena &amp; Pickering, 2021; Liang et al., 2021; Sangkaew &amp; Zhu, 2020; Collado et al., 2020; Marine-Roig, 2019; De Maio, Forsini, Messina, Santoro &amp; Violi, 2020; Milwood &amp; Crick, 2021.</td>
</tr>
</tbody>
</table>

Source: Authors (2021)

Cluster 3 refers to several tools and methodologies that are being used on the internet related to SA. First, it is important to highlight that different authors search to comprehend the employ of SA and how this methodology impacts tourism research (Li, Zhu, Guo & Cambria. 2020a; Moreno-Ortiz; Salles-Bernal & Aorta-Varea, 2019; Vázquez Loaiza, Pérez-Torres & Conteras, 2019; Fu, Hao, Li & Hsu, 2019; Li, Guo, Shi & Zheng, 2018b; Kirilenko et al., 2018; Liu, Tian, Feng & Zhuang, 2018; González, Cámera, Valdivia & Zafra, 2015; García et al., 2012); on hotel management (Liang, Liu & Wang, 2019; García-Pablos et al., 2016) and on management (Angskun & Angskun, 2019; Wang, Chiang & Sun, 2019; Yang & Chao, 2018; Marrese-Taylor et al., 2018).

Table 3: Main topic and authors - Cluster 03

<table>
<thead>
<tr>
<th>Cluster 03</th>
<th>Articles in this cluster discuss the several tools and methodologies that can be used to use the corpus used to make SA.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Francisco e Cristina Manosso &amp; Thays Cristina Domareski Ruiz</td>
<td>Li et al., 2020a; Moreno-Ortiz et al., 2019; Vázquez Loaiza et al., 2019; Fu et al., 2019; Angskun &amp; Angskun, 2019; Wang et al., 2019; Yang &amp; Chao, 2018; Marrese-Taylor et al., 2018; Li et al., 2018b; Li et al., 2018c; Kirilenko et al., 2018; Li et al., 2018d; Collado et al., 2015; García et al., 2012.</td>
</tr>
</tbody>
</table>

Source: Authors (2021)

Finally, Cluster 4 complements the discussions in Cluster 2, bringing new perspectives to the investigation about tourism and SA. Here, the articles investigate destination knowledge (Antonio, Correia & Ribeiro, 2020); destination branding (Otay Demir, Yavuz Gökmen & Rafillery, 2021; Chen, Liu, Wang & Chen, 2019); racism in online reviews (Li, Li, Law & Paradies, 2020b); cruise and destination image (Wu et al., 2020); destination crises (Gkritzali, Mavragani & Grätzilis, 2019); and the role of Twitter to understand sentiments and consumer behavior (Papapicco & Mininini, 2020; Bolic, Acciarini, Marchegiani & Pirol, 2020; Feizollah, Ainin, Anuar, Abdullah & Hazim, 2019).

Table 4: Main topic and authors - Cluster 04

<table>
<thead>
<tr>
<th>Cluster 04</th>
<th>It is a complementation of Cluster 02 and brings new perspectives to the investigation about the relations between SA and tourism research.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Francisco e Cristina Manosso &amp; Thays Cristina Domareski Ruiz</td>
<td>Otay Demir et al., 2021; Papapicco &amp; Mininini, 2020; Bolic et al., 2020; Antonio et al., 2020; Li et al., 2020b; Wu et al., 2020; Chen et al., 2019; Grätzilis et al., 2019; Feizollah et al., 2019.</td>
</tr>
</tbody>
</table>

Source: Authors (2021)

It is possible to observe that clusters 1, 2, and 4 are more related to tourism practice than cluster 3, which emphasizes the application of different methodologies that can be used. In Table 5, we present the four clusters and their keywords. Therefore, it is highlighted that the articles discussed in one cluster may be part of another, as the discussion about SA and tourism covers different perspectives of the four clusters.

Table 5: Clusters and keywords

<table>
<thead>
<tr>
<th>Clusters</th>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 01</td>
<td>Sentiment Analysis (SA) Online Reviews (OR) Machine Learning (ML)</td>
</tr>
<tr>
<td>Cluster 02</td>
<td>Social Media (SM) Tourism Destination (TD) Social Networking (SN) Sentiment Mining (SM)</td>
</tr>
<tr>
<td>Cluster 03</td>
<td>Data Mining (DM) Social Networking (SN) Opinions Mining (OM) Drop-earning (DE)</td>
</tr>
<tr>
<td>Cluster 04</td>
<td>Tourism (TR) Twitter (TW) Internet (I) Article (A) Human (H)</td>
</tr>
</tbody>
</table>

Source: Authors (2021) based on Vosviewer data

In addition, it is relevant to present the most cited articles that are considered empirical and theoretical studies related to tourism research that employ SA. In Table 2, for instance, the methodological approach of the top-cited articles mixed quantitative and qualitative tools, which means that when the researchers use SA as a way to understand what emotions arise from a tourist experience, it is possible to congregate or, even, choose the methodological approach that is most convenient for the study.

Table 6: Profile of Typical Travel and Tourism Research

<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Year</th>
<th>Methodological Approach</th>
<th>Citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>What do Airbnb users care about? The analysis of online review comments</td>
<td>Cheng &amp; Ni</td>
<td>2019</td>
<td>Qualitative</td>
<td>116</td>
</tr>
<tr>
<td>Semi-supervised learning for big social data analysis</td>
<td>Hassanein &amp; Cartwright</td>
<td>2018</td>
<td>Quantitative, bIVM and HLS</td>
<td>92</td>
</tr>
<tr>
<td>Predicting hotel review helpfulness: The impact of review visibility, and interaction between hotel scores and review range</td>
<td>Hu &amp; Chen</td>
<td>2018</td>
<td>Quantitative, MedePLOS</td>
<td>78</td>
</tr>
<tr>
<td>Automated Sentiment Analysis in Tourism: Comparison of Approaches</td>
<td>Kirilenko et al.</td>
<td>2018</td>
<td>Quantitative, Sentiment Analysis</td>
<td>90</td>
</tr>
<tr>
<td>The influence of tourists’ emotions on the selection of particularly word of mouth platforms</td>
<td>Yang et al.</td>
<td>2018</td>
<td>Quantitative, Correlation analysis</td>
<td>36</td>
</tr>
<tr>
<td>Monitoring the environment and human sentiment on the Great Barrier Reef Using sentiment analysis</td>
<td>Becken et al.</td>
<td>2017</td>
<td>Qualitative</td>
<td>34</td>
</tr>
<tr>
<td>“I will never go to Hong Kong.” How the secondary crisis communication “Occupied Central” on Weibo affects tourists to a tourist boycot</td>
<td>Lao &amp; Chen</td>
<td>2017</td>
<td>Qualitative and Quantitative, ROST, Content, using System 6.0</td>
<td>34</td>
</tr>
<tr>
<td>Learning to the voices from home: An analysis of Chinese tourists’ sentiments regarding Australian destinations</td>
<td>Liu et al.</td>
<td>2019</td>
<td>Qualitative, Techniques of lexicon filtering, Co-occurrence analysis, Sentiment mining, The model analyzes online reviews, Sentiment image analysis</td>
<td>31</td>
</tr>
</tbody>
</table>

Source: Authors (2021) based on Scopus database
These studies are considered innovative, by presenting several techniques of qualitative and quantitative methodologies. In this case, tourism appears to be an ideal application field of social media analytics with tremendous growth and potential (Xiang; Du; Ma; Fan, 2017; Gretzel et al., 2012) where SA can provide some interesting insights related to destination management, consumer behavior, past experiences, satisfaction, quality of services. All these topics can be discussed and analyzed by a SA approach and further empirical investigations can use this tool to comprehend them and in consequence provide structuralized information not only to academics but also to entrepreneurs and who are responsible for public policies.

It is possible to observe that the most cited articles present qualitative (3), quantitative (6), and qualitative and quantitative analysis (1). This type of analysis allows an integrated approach, able to detect qualitative and/or quantitative factors. To exploit the data, it is necessary to adopt rigorous methodologies consistent with ICTs and big data. SA can rely on the application of established techniques, advanced by the diffusion of software for data processing and analysis.

The application of algorithms and procedures in an increasing number of fields has favored diverse and specialized contributions and studies, thanks only to their application to a wide range of phenomena and processes, essentially within social networks (Micera & Crispino, 2017). All the papers are using SA with online reviews, analyzing accommodation, marketing, consumer behavior, or experience.

5 DISCUSSION

SA has as the object of research a review or commentary about a product or service that has been made public on the Internet domain (Mäntyla et al., 2018). As the current research shows, there is a growing interest in comprehending consumer behavior through their opinions on the different social networks, mainly, on the tourism domain. In this topic, an integrative discussion is made, which means that subsequent to the bibliometric analysis in the Vosviewer software, an analysis of the main subjects was engendered and discussed more profoundly.

First, it is relevant to highlight that, according to Alaei, Becken, and Stantic (2017), tourism is an industry where the customer experience is essential for its growth and reputation, using big data and new data sources became a necessity to understand the customer needs, feelings, and desires. Consequently, SA was transformed into a new methodology to decode the online user-generated content available and to understand perceptions and characteristics of different market groups (Sigala et al., 2002; Ribeiro et al., 2016; Christou et al., 2021).

In tourism research, the discussion about SA has become more prominent in recent years. Hence, the present study seeks to discern which themes have been linked with SA as a method to understand consumer behavior in the tourism domain. Here, it is important to highlight that some research discusses how to use SA as a tool and not the topics that have been discussed in tourism research. Consequently, the discussion proposed in this topic debates this gap. Figure 5 shows the themes that are relevant to the debate that is divided into five areas, such as management, information, and communication technologies (ICT); methodology; marketing, and tourism sectors.

Figure 5: Discussion Areas in Tourism Research and Sentiment

Management is an area that has been used SA to understand the consumer necessities, desires and, in consequence, draw new products and services for them. In this topic, it is possible to observe that tourism, in general, has 16 papers that discuss SA and its use from the management perspective. For instance, Antonio et al. (2020) present topics that could be used by different destination management organizations to promote these cities and even have better destination knowledge from the content that tourists have been sharing in the social networks. Also, De Maio et al. (2020) presents a project that uses social data to draw a tour planner with optimized itineraries based on the user's profile; in turn, Zapata et al. (2019) describe how to use SA to design new business information architecture to allow touristic enterprises make better decisions. Micera and Crispino (2017) and Kim et al. (2017) uses SA as a methodology with the context of smart destinations to analyze the perceptions of the tourists; this analysis suggests the use of this information in the marketing and management planning process of the destinations.

Becken et al. (2019), in the context of tourism management, apply SA to understand through Twitter posts the emotions that arise from a visit to a destination and how this could be used to deliver better experiences to the tourist. In complementation to this perspective, Liu et al. (2019) discuss the sentiments that emerge from a visit to different Australian destinations, using the Chinese tourists for this matter and, Starosta, Budz, and Krutwig (2019) utilize the german speakers to comprehend the experiences in European destinations. Another issue discussed in the management concern is the fake reviews and their impacts on the tourism business (Reyes-Menendez, Saura & Filipe, 2019).

Besides the tourism perspective, some research discusses the management angle in the context of national parks and geoparks (02 papers), with Mangachena and Pickering (2021) talking about how social media can be employed in a new
way to manage national parks through SA and Twitter posts. Finally, the cultural perspective also has been considered in the management context (04 papers) with the halal tourism discussion (Ainin et al., 2020; Feizollah et al., 2019) and the cultural perceptions in the tourist travel notes to the destination management (Liang et al., 2021).

Information and communication technology (ICT) is another perspective that has been discussed in tourism research, big data, for example, has 07 published papers and bring out the following considerations: an extensive data analysis to comprehend the impact of cruise tourism image and how its impacts their satisfaction and purchase intentions (Wu et al., 2020); Önder, Gunder, and Scharl (2019) debates about forecasting tourist arrivals with web sentiment, which means using big data and new media to understand the consumer behavior; Imane and Abdelouahab (2019) employs social big data to comprehend the guest's experiences in a Five Stars hotel; in turn, Hussain and Cambria (2018) highlights the insertion of big social data in the tourism research and how it could be done in different approaches. In addition to this perspective, Bolici et al. (2020) argue about innovation diffusion in tourism, analyzing how information about blockchain is spread in Twitter posts. It is possible to highlight, in this case, that ICT has been used SA as a way to understand consumer behavior from big data analytics and the impacts of innovation diffusion in the context.

The methodological approach, named SA and its derivatives were also presented in several analyzed articles, such as Luo, Zhang, Qin, Yang & Liang (2021b), which discusses the Probabilistic Linguistic Term and IDOCRIW-COCOSO Model; Li et al. (2020a) highlight the employment of lexicon integrated two-channel CNN-LSTM family models to verified SA in the user reviews; Shi, Zhu, Li, Guo & Zhang (2019) argues about classic and the latest textual SA approach and its use in different sectors, as tourism; Wang et al. (2019) brings out the multi-lexicons use in the context of touring reviews. In this study area it is relevant to highlight that each article analyzed brings a new perspective to SA and opinion mining approach, which means that each author presents a different perspective inside this methodology, but as mentioned, some of them discuss more profoundly the method and its impact in the tourism industry.

Marketing has been used SA as a way to understand consumer behavior and also draws news products, services, and experiences. Hence, in the tourism context, it is not different because the articles analyzed discuss consumer behavior (17), online reviews (12), destination image (5), storytelling (3), and branding (2). Applying SA in marketing destinations could reduce research costs and time (Kim et al., 2017).

SA has as its objective to understand the consumers' feelings and emotions and, in some cases, develop new experiences that make the tourist, in the case of the present research, more satisfied and with the intent to recommend and return in the hotel, restaurant and, even, in the destination. Thus, it is possible to discuss the consumer behavior and online reviews perspective jointly because both searches to comprehend how products and services impact the overall experience. Marcolin et al. (2021) and Nave, Rita, and Guerreiro (2018) present a new perspective using online reviews to improve decision-making processes, mainly in the context of managers and the necessity to draw experiences that are positive to their customers. Aggarwal and Gour (2020) debate the relevance to peek inside the consumers' minds, searching their opinions and debating about how it can be employed. For its parts, Li et al. (2020b) highlights the racism in the online reviews and discussions about racial discrimination on tourist's experience, an important question to comprehend this context in the destination management and, even, in the consumer behavior perspective.

Online reviews and their relevance in the construction of new experiences for tourists is discussed by Sun et al. (2020); Lee et al. (2020); Sangkaew and Zhu (2020), Vallone and Veglio (2019), and others that use the construct satisfaction and search in the online reviews how it can be used to manage the creation of new services that are related to the needs and consumers' desires. Therefore, the consumers' perspective is discussed in those papers with the consumer engagement and the role of online reviews (Kesgin & Murthy, 2019; Chatzigeorgiou & Christou, 2020), and the relevance of the SA to build qualified information (Yang and Chao, 2018). Besides this context, the discussion about destination image is made by Marine-Roig (2019); Gkritzali et al. (2018); Micera and Crispino (2017), and others that apply SA as a way to manage the destination image. At least, storytelling, mainly the digital perspective, has being used the SA approach to improve their content and tell better stories for the consumer (Zhang, Kim, Kim & Fesenmaier, 2019; Zhang, Choe & Fasenmaier, 2019; Zhang & Fasenmaier, 2018).

The tourism sector in the analyzed papers brings up debates about hotels (14), food and beverage (01), transportation (01), among others. For the present discussion, it is important to highlight how SA has been used in the hotel's perspective, the highest theme in the discussion. For instance, Kim et al. (2021) points out that visual information impacts the consumer responses in online reviews and needs to be better understood to influence the design experience positively. In addition, An et al. (2020) discusses user-generated photos as a new way to comprehend consumer and tourist behavior. Ray et al. (2021) debated a way to categorize the hotel reviews and draw a recommender system with this content; Zhu, Lin, and Cheng (2020) and Park et al. (2020) discuss the construct satisfaction and trust in online reviews and how its impact the way the consumer sees the hotel and if it is a key to make him revisit.

Moreover, the papers reflect about SA and other questions, like the environmental issue with the discussion about air pollution (Zhang, Yang, Zhang & Zhang, 2020; Tao, Zhang, Shi & Chen, 2019); COVID-19 (Sontasayara et al., 2021) and other relevant questions that can be better understood, because there are only a few numbers of articles about those issues. In this case, it is meaningful to assimilate the main topics that the SA approach has been utilized.

In the last decades, society has been faced a technological revolution, which means that the consumer (i.e., the tourist) use the content and the information shared online to draw their own experience and also, present to the entrepreneurs their necessities, desires, dissatisfaction, and other information that can be considered as essential for the market development (Inversini et al., 2015). Thus, SA became one of the several tools to comprehend consumer behavior, employed as a way to look at the feelings (i.e., positive,
negative, or neutral), emotions, and opinions about certain products, services, and experiences (Kirilenko et al., 2018). Consequently, our findings make a studies’ compilation that discusses SA, and it is relevant, because not only researchers can use this data, but also people who are in the market. The papers presented discuss several implications for the debate about SA and tourism, which means that when we observe the research corpus, it is possible to comprehend the uses and benefits of this tool. As mentioned, previous research did not compile the subjects and analyzed more profoundly the themes and concerns around SA and tourism. So, the present research looks to fill this gap and construct a broad vision on the subject.

When we analyze the academic implications of our findings, it is important to highlight that a bibliometric, systematic, and integrative review helps the researchers to understand the field that they propose to study. In the case of SA and tourism, the main findings suggest the major topics that have been used this methodology as a way to understand the massive information provided by the new technologies and the content shared by the consumer on several platforms (Micera and Crispino, 2017). Besides, researchers can use this data to justify the SA approach in a specific field, because the results help them to observe what are the gaps and the discussions that need more attention and discussions. In addition, it is essential to debate the management implication of our study. First, when an entrepreneur reads this kind of research, they can have a broad perspective about a subject, in this case, SA and tourism. Also, it is possible to observe how this method has been used in several domains. It enables the application of the tool in the organizational context, which means that the results present for the entrepreneurs and managers new paths within that field.

Further, demonstrate the importance of using certain instruments (e.g., SA) to build increasingly competitive and relevant businesses in the market. Digital communication has been used to reach many people in a short time, which means that when a destination, hotel, restaurant, and other tourist facilities employ this kind of mechanism to create a bond with their consumers. Much information can be produced in this relationship. It happens because a new way of interaction emerges, and every consumer can share their opinions online, leading to a necessity for a major comprehension of what has been disseminated on different platforms (Todisco et al., 2020).

As seen in the results, it is possible to demonstrate that SA is a methodology that has gained prominence in the last years (Christou, 2010; Kim et al., 2017; Bonarou, 2021). It happens because understanding the sentiments, emotions, and feelings expressed by consumers in their opinions and stories shared on social media became relevant content to draw products, services, and experiences that make tourists more satisfied with their past purchases and experiences. Finally, SA and opinion mining became tools to gather the information and show the managers how their consumers feel about what has been offered in the actual market and if changes are necessary to deliver better experiences. In this case, from an academic perspective, researchers can use this information to discuss the subject in a scientific and theoretical context, in other words, sentiment analysis is used as a methodological tool for several projects in the tourism field and as a way to put together people that are related to businesses and academics.

6 CONCLUSIONS

The rapid diffusion of the use of the internet has dramatically changed the habits of people who spend more time online, creating interconnected networks impacting the tourism sector (Micera & Crispino, 2017). Lately, with the impact of ITCs, the demand for big data analysis has increased in the hospitality field. This study analyzes how SA has been used in tourism research. This bibliometric study is beneficial not only for researchers but also for decision-makers in public and private organizations because the findings are relevant. It is possible to show the government and entrepreneurship how to draw communication and marketing strategies, contributing to the decision-making process, and comprehend consumer behavior. Literature reviews help to consolidate and advance theory providing insights to address scholars in their efforts (Del Chiappa et al., 2021; Fotiadis, 2018; Fotiadis & Williams, 2018).

As the main results, it is possible to analyze four clusters that appear in the context of SA and tourism, and each one discusses a different perspective to the use of this technique in tourism studies. It reveals the article characteristics, top citations and geographical distribution, the journals that publish the topic, and the co-occurrence of author keywords, highlighting 4 clusters that are related. The main topics that are using SA in tourism are marketing (consumer behavior and online reviews), methodology (opinion mining tools), ICTs (big data), management (tourism), tourism sector (hotels).

SA can represent the degree of positivity or negativity of the data but has little prescriptive and practical implications. However, using SA, we can find out how tourists perceive negatively or positively about certain destination services but cannot discover why they feel like that.

As a theoretical implication, both hospitality and tourism management cases based on online reviews, big data analysis, SA were successfully combined to derive more meaningful research outcomes that can provide interesting insight to be applied in the tourism sector. The limitations of this study should be recognized. The findings could be extended by using other well-known databases such as Web of Science and considering other sources of information, such as articles in different languages other than English. As suggestions for future studies, other areas in the tourism sector could be more explored as food and beverage, transportation, and communication crisis that have been identified in this research. The use of other software that could run more extensive samples is also an alternative to enhance the studies with SA.

REFERENCES

Carroll, C. D., Cambria, E., Schuller, B., Xia, Y., Havasi, C. (2013). New avenues...
mining model based on machine learning approach. Tourism Management Perspectives, 37, 100769.


